

# MUSIC RETRIEVING DEVICE AND METHOD, AND STORAGE MEDIUM RECORDED WITH MUSIC RETRIEVING PROGRAM

Publication number: JP2000347659 (A)

Also published as:

Publication date: 2000-12-15

JP3631660 (B2)

Inventor(s): NISHIHARA YUICHI; KOSUGI NAKKO; SAKATA TETSUO;  
YAMAMURO MASASHI; UMEDA MASAYOSHI; KONYA SEIICHI  
\*

Applicant(s): NIPPON TELEGRAPH & TELEPHONE +

Classification:

- international: G06F17/30; G10H1/00; G10H1/40; (IPC1-7): G06F17/30;  
G10H1/00; G10H1/40

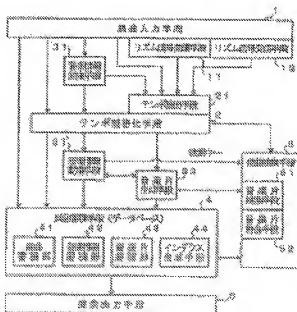
- European:

Application number: JP20000008458 20000118

Priority number(s): JP20000008458 20000118; JP19990082722 19990326

## Abstract of JP 2000347659 (A)

**PROBLEM TO BE SOLVED:** To precisely and quickly retrieve music intended by a retriever out of a database, by inputting freely a portion or the whole of the music without making the retriever be aware of a phrase, a melody, a tempo and the like.  
**SOLUTION:** A tempo of input music or musical information converted from the input music is normalized by a tempo normalizing means 2 to be stored in a database by a music accumulating means 4, in database construction. The tempo of the music or the musical information converted from the music, which is a retrieving key, is normalized by the tempo normalizing means 2, similarity retrieval is carried out by a music retrieving means 5 based thereon, and the music as a retrieved result is output by a music outputting means 6, when retrieved.



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2000-347659

(P2000-347659A)

(43) 公開日 平成12年12月18日 (2000.12.18)

(51) Int.Cl. <sup>7</sup>	識別記号	P I	サーチコード (参考)
G 1 0 H 1/00	1 0 2	G 1 0 H 1/00	1 0 2 Z
G 0 6 F 17/30		G 0 6 F 1/40	
G 1 0 H 1/40		G 0 6 F 15/40	3 7 0 E
		15/403	3 5 0 C

審査請求 未請求 請求項の数 14 O L (全 12 頁)

(21) 出願番号	特願2000-8458 (P2000-8458)	(71) 出願人	00004226 日本電信電話株式会社 東京都千代田区大手町二丁目3番1号
(22) 出願日	平成12年1月18日 (2000.1.18)	(72) 発明者	西原 祐一 東京都千代田区大手町二丁目3番1号 日 本電信電話株式会社内
(31) 優先権主張番号	特願平11-82722	(72) 発明者	小杉 尚子 東京都千代田区大手町二丁目3番1号 日 本電信電話株式会社内
(32) 優先日	平成11年3月26日 (1999.3.26)	(74) 代理人	100087548 菅野士 小笠原 吉壽 (外1名)
(33) 優先権主張国	日本 (J P)		

最終頁に続く

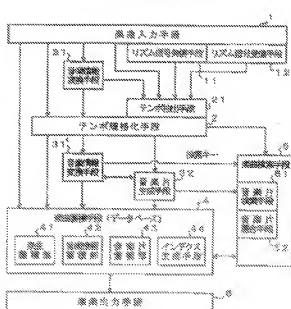
(54) 【発明の名称】 音楽検索装置、音楽検索方法および音楽検索プログラムを記録した記録媒体

(57) 【要約】

【課題】 検索者が、フレーズ、調、テンポ等を意識することなく、自由に楽曲の一部または全体を入力することにより、検索者が登録した楽曲をデータベース中から精度よく、かつ高速に検索できるようにする。

【解決手段】 データベースの構成では、入力された楽曲もしくは楽曲から変換された音楽情報のテンポを、テンポ規格化手段2により規格化し、楽曲登録手段4によりデータベースに格納する。検索の際には、テンポ規格化手段2により検索キーとなる楽曲もしくは楽曲から変換された音楽情報のテンポを規格化し、それをもとに楽曲検索手段3により類似検索を行い、楽曲出力手段6により検索結果の楽曲を出力する。

本発明の構成構成図



## 【特許請求の範囲】

【請求項1】 楽曲を検索する検索装置において、楽曲を入力する楽曲入力手段と、上記楽曲入力手段により入力された楽曲のテンポを規格化するテンポ規格化手段と、少なくとも上記テンポ規格化手段によりテンポ規格化された楽曲を蓄積する楽曲蓄積手段と、検索キーとして入力された楽曲の一部または全体から上記テンポ規格化手段によりテンポ規格化された楽曲の一部または全体をもとに、上記楽曲蓄積手段に蓄積された楽曲の一部または全体の中から類似した楽曲を検索する楽曲検索手段と、上記楽曲検索手段により検索された楽曲を出力する楽曲出力手段とを有することを特徴とする音楽検索装置。

【請求項2】 請求項1記載の音楽検索装置において、上記楽曲入力手段により入力された楽曲を音楽情報に変換する音楽情報変換手段を有し、上記テンポ規格化手段は、楽曲を上記音楽情報変換手段により音楽情報に変換した後でテンポ規格化を行うことを特徴とする音楽検索装置。

【請求項3】 請求項1または請求項2記載の音楽検索装置において、上記楽曲入力手段は、一定の時間間隔もしくは決められた規則に従った時間間隔で信号を発信するリズム信号発信手段、または、楽曲入力者が入力中の楽曲のリズムを合わせて入力できるようなリズム信号受信手段を有することを特徴とする音楽検索装置。

【請求項4】 請求項1ないし請求項3のいずれかに記載の音楽検索装置において、上記テンポ規格化手段は、上記楽曲入力手段により入力された楽曲、上記音楽情報変換手段により変換された音楽情報、上記リズム信号発信手段により発信された信号、上記リズム信号受信手段により受信された信号の少なくとも一つを用いてテンポを抽出するテンポ抽出手段を有し、抽出されたテンポを用いてテンポ規格化を行うことを特徴とする音楽検索装置。

【請求項5】 請求項1ないし請求項4のいずれかに記載の音楽検索装置において、上記楽曲検索手段は、テンポ規格化された検索キーの一部または全体をもとに、上記楽曲蓄積手段に蓄積された、テンポ規格化された楽曲の一部または全体を検索する際に、テンポ規格化された検索キーの一部または全体と、テンポ規格化された検索キーの一部または全体と同一長さの、検索対象となるテンポ規格化された楽曲の一部または全体との類似度を算出することを特徴とする音楽検索装置。

【請求項6】 請求項1ないし請求項5のいずれかに記載の音楽検索装置において、上記テンポ規格化された楽曲もしくはテンポ規格化された音楽情報に対して、音楽データの分割を行う音楽片生成手段を有することを特徴とする音楽検索装置。

【請求項7】 請求項6記載の音楽検索装置において、上記音楽片生成手段は、音楽片を生成する際にすべての

音楽片の長さを一定にすることを特徴とする音楽検索装置。

【請求項8】 請求項6または請求項7記載の音楽検索装置において、上記音楽片生成手段は、音楽片を生成する際に、同じ楽曲に対する音楽片は、一定の時間間隔で音楽片が開始するように生成する、あるいは音楽片が音符の始まりから開始する、または音の高さの変動もしくは変動率がある一定の閾値を超えた点から開始するように生成することを特徴とする音楽検索装置。

【請求項9】 請求項1ないし請求項8のいずれかに記載の音楽検索装置において、上記楽曲蓄積手段に蓄積する前に音楽情報への変換を行う音楽情報変換手段を有することを特徴とする音楽検索装置。

【請求項10】 請求項1ないし請求項9のいずれかに記載の音楽検索装置において、楽曲を蓄積する際に、検索を高速化するためのインデックスを生成するインデックス生成手段を有し、上記楽曲検索手段は、検索を行う際にインデックスを用いることを特徴とする音楽検索装置。

【請求項11】 請求項6ないし請求項10のいずれかに記載の音楽検索装置において、上記楽曲検索手段は、検索を行う際に、音楽片同士を検索を行い、その結果の音楽片同士の関係から最終的な検索結果を導く音楽片検索手段を有することを特徴とする音楽検索装置。

【請求項12】 請求項11記載の音楽検索装置において、上記楽曲検索手段は、検索キーの音楽片同士の関係と、検索結果の音楽片同士の関係の整合性を調べた後、最終的な検索結果を導く音楽片整合手段を有することを特徴とする音楽検索装置。

【請求項13】 楽曲を検索する検索方法において、楽曲を入力する過程と、入力された楽曲のテンポを規格化する過程と、少なくとも、テンポ規格化された楽曲を蓄積する過程と、検索キーとして入力された楽曲の一部または全体から得られた、テンポ規格化された楽曲の一部または全体をもとに、類似した楽曲の一部または全体の中から類似した楽曲を検索する過程と、検索された楽曲を出力する過程とを有することを特徴とする音楽検索方法。

【請求項14】 計算機によって楽曲を検索するために用いるプログラムを記録したプログラム記録媒体であって、楽曲を入力する処理と、入力された楽曲のテンポを規格化する処理と、検索キーとして入力された楽曲の一部または全体から得られた、テンポ規格化された楽曲の一部または全体をもとに、あらかじめテンポ規格化された楽曲の一部または全体となる楽曲の一部または全体の中から、類似した楽曲を検索する処理と、検索された楽曲を出力する処理とを、計算機に実行させるプログラムを記録したことを特徴とする音楽検索プログラムを記録した記録媒体。

【符号の説明】

【0001】

【発明の属する技術分野】本発明は、音楽検索に係わり、特に、楽曲入力手段によって入力もしくは指定された楽曲の一部もしくは全体を検索キーとして、データベースに蓄積された楽曲の中から、検索キーと類似する楽曲の一部もしくは全体を検索する装置およびその方法に関する。

#### 【0002】

【従来の技術】音楽の類似検索を行う場合、従来方法においては音符がマッチングの処理単位として用いられていた。この場合、音符とは、音の高さの変動もしくは変換率がある一定の閾値内に納まっている音の連続のことを言う。また、各音符は、音符情報として音の高さ（音高）、音の長さ（音長）、時間的に前後する音符との音高の差（相対音高）、時間的に前後する音符との音長の比（相対音長）などを保有している。

【0003】図1は、従来方法による音楽検索のうち、音符を単位としたDIPマッチングを使用するものの原理を説明する図である。まず、データベース構築のフェーズでは、検索対象とする楽曲を入力し（ステップS101）、音符情報に変換する（ステップS102）。それをデータベースへ格納する（ステップS103）。データベースへ格納した楽曲の検索は、検索キーとなる楽曲を入力し（ステップS111）、それを音符情報に変換する（ステップS112）。次に、データベース中の全楽曲の音符情報と検索キーから得られた音符情報について、データベース中の曲数分、DP（Dynamic Programming）マッチングを実施し（ステップS113）、もっとも類似した楽曲の候補を検索結果として出力する（ステップS114）。

【0004】以上のように、音楽の類似検索を行う場合、従来方法においてはDIPマッチングが使われていた（参考文献1、2）。

（参考文献1）田田智也、後藤真幸、村岡洋一：「WWW上での音声による曲検索システム」、電子情報通信学会技術報告、SP97-103、1998。

（参考文献2）嶋山哲也、高橋幸典：「ハミング歌を手掛りするメモリ検索」、電子情報通信学会論文誌、Vol. J77-D-11 No. 8 pp. 1543-1551、1994。

DIPマッチングにおいては、検索キーと、データベース中の楽曲の一部または全体（候補部分曲）の対応関係が求められ、対応関係にある検索キーの音符と、候補部分曲の音符との間で、音符情報の差の総計によって類似度が計算される。この検索キーとの間の類似度が最大になるような候補部分曲を、最終的検索結果として出力するのがDIPマッチングである。

#### 【0005】

【発明が解決しようとする課題】従来方法の音符をマッチングの単位とした検索方法では、一般に、音符ごとの類似度の差を計算している。

【0006】また、従来方法のDIPマッチングでは、計

算量を抑制するために、検索キーの音符列と、類似度計算対象となるデータベース中の楽曲の音符列との間の音符数の差に制限を設ける。

【0007】したがって、音楽検索をする場合に、検索キー入力の際に、検索目標とするデータベース中の楽曲に対して一定量以上に音符の挿入・削除を行ってしまうと極端に検索精度が悪くなるという問題があった。

【0008】たとえば検索キー入力の際に、音符の挿入・削除が行われなかったとしても、検索キーから音符情報を抽出する際に、音符の挿入・削除が生じる可能性も大きく、そのために検索精度が悪いという問題があった。

【0009】また、DIPマッチングにおいては、インデクスを利用することができず、データベース中の全曲に対して全検索を施す必要があるために、たとえ計算量を抑制するために許さざるを得る音符数減りの量を制限したとしても、やはり検索に時間がかかるという問題があった。

【0010】一方、音楽片を使用した検索の場合、インデクスを利用してするために検索時間を大幅に短縮することが可能であった。

【0011】図1は、音楽片を使用する音楽検索の原理を説明する図である。この手法は、特願平10-341754号（音楽検索装置、音楽検索方法および音楽検索プログラムを記録した記録媒体）で、本発明者等が提案しているものである。

【0012】まず、データベースの構築では、検索対象とする楽曲を入力し（ステップS201）、入力した楽曲を音符情報に変換し（ステップS202）、その音符情報を複数の音楽片に分割し（ステップS203）、前述検索が可能のように、インデクスを生成し（ステップS204）、音楽片の情報をインデクスとともにデータベースに格納する（ステップS205）。

【0013】次に、楽曲の検索では、検索キーとなる楽曲を入力し（ステップS211）、入力した楽曲を音符情報に変換し（ステップS212）、音符情報を複数の音楽片に分割する（ステップS213）。分割片のそれぞれを検索キーとして、インデクスを用いてデータベースに蓄積された音楽片から類似するものを検索する（ステップS214）。検索結果として抽出された音楽片同士の間係と、検索キーとなった音楽片同士の間係などを調べることによって、抽出された音楽片を含むデータベース中の楽曲の一部または全体との間の類似度を調べて（ステップS215）、検索キーとなった楽曲の一部または全体と類似度の高いデータベース中の楽曲の一部または全体を検索結果として出力する（ステップS216）。

【0014】音楽片検索を行う場合には、類似度を計算する音楽片同士の間係の値が一致しているために、DIPマッチングを用いて音符同士の類似度を求める必要がなく、インデクスを利用した検索が効率的に可能であった。

【0015】しかし、音楽片検索をする際は、音符数部りが認められないために、音楽片の長さ分だけ、音符の挿入・削除のない空間が必要であるという制約があった。すなわち、音楽片の長さ分だけ、音符の挿入・削除のない空間がなければ、検索精度が悪いという問題点があった。

【0016】また、音符情報を含むしも必要としない方法としては、音楽情報への変換を行った後、類似検索を行うという手法がある。

【0017】図11は、音符情報を使用しない音楽検索の原理を説明する図である。この手法は、特願平10-329131号(音楽情報検索装置、音楽情報蓄積装置、音楽情報検索方法、音楽情報蓄積方法およびそれらのプログラムを記録した記録媒体)で、本発明者が提案しているものである。

【0018】まず、データベースの構築において、楽曲を入力し(ステップS301)、入力した楽曲から音楽的特徴(特徴量)を抽出することが可能な音楽情報に変換し(ステップS302)、それをデータベースへ格納する(ステップS303)。

【0019】楽曲の検索において、検索キーとなる楽曲を入力し(ステップS311)、入力した楽曲を同等な音楽情報への変換を行い(ステップS312)、変換した音楽情報をもとに、データベース中の楽曲について類似検索を実行する(ステップS313)。ここで、類似検索として、検索キーの楽曲の特徴量をデータベースに登録された楽曲の特徴量と比較し検索し、検索は距離関数を用いて距離の近いものを類似度の高いものとする類似度計算を行う。類似したものを含む楽曲の検索結果を出力する(ステップS314)。

【0020】しかし、このような音楽情報を利用する方法においては、検索キーと、データベース中の曲とのテンポが異なる場合に、検索精度が悪いという問題点があった。

【0021】本発明は、上記の点に鑑みながら、データベース中の楽曲および検索キーのテンポを規格化することにより、音符を単位とするのではなく、規格化された時間を単位とした検索を可能とすることによって音符数部りによって影響を受けない楽曲の検索精度をより向上させることを目的とする。本発明において、ロドマッチングを行う必要があるため、インデックスを用いた検索を行うことが可能であり、高度な検索が可能である。また、検索を行う前にテンポを規格化することにより、検索キーとデータベース中の曲とのテンポが異なる場合にも、高い検索精度を得ることが可能である。

【0022】

(実施例) 本発明を説明するための手段(図1および図2)は、本発明による音楽の類似検索の概要を説明するための図、図3は、本発明の処理構成図である。

【0023】本発明は、音楽検索において、データベースに入力された楽曲もしくは楽曲から変換された音楽情報のテンポを規格化した上でデータベースに蓄積し、また、検索キーとして入力された楽曲もしくは楽曲から変換された音楽情報のテンポを規格化した上で、音楽の類似検索を行うことを主な特徴とする。

【0024】まず、音楽情報への変換を行わない検索におけるデータベースの構築では、楽曲入力手段1を用いて楽曲が入力される(ステップS10)。入力された楽曲は、テンポ規格化手段2によってテンポ規格化される(ステップS11)。

【0025】また、音楽情報への変換を行う検索の場合のデータベースの構築では、楽曲入力手段1を用いて楽曲が入力された後(ステップS30)、音楽情報変換手段3により音楽情報に変換され(ステップS31)、その後、テンポ規格化手段2によってテンポ規格化される(ステップS32)。

【0026】テンポ規格化された楽曲あるいはテンポ規格化された音楽情報は、楽曲蓄積手段4により、音楽情報蓄積部42にデータベースとして蓄積される(ステップS12、S33)。

【0027】テンポ規格化された楽曲あるいはテンポ規格化された音楽情報は、楽曲蓄積手段4により蓄積される前に、音楽情報変換手段3'による音楽情報への変換、音楽片生成手段32による音楽片の分割の処理を経てから、蓄積されることもある。さらに、蓄積された楽曲、音楽情報、音楽片に対して、インデックス生成手段44により、検索を高速化するためのインデックスが作成されることもある。

【0028】楽曲蓄積手段4には、テンポ規格化される前の楽曲を楽曲蓄積部41に蓄積することも可能であり、検索の結果として、テンポ規格化される前の楽曲を出力することも可能である。

【0029】楽曲の検索においては、まず、楽曲入力手段1を用いて楽曲の一部または全体が入力もしくは指定される(ステップS20、S40)。入力された楽曲は、テンポ規格化手段2によりテンポ規格化される(ステップS21)。音楽情報変換手段31により音楽情報へ変換された後(ステップS41)、テンポ規格化される(ステップS42)。テンポ規格化された検索キーは、さらに、音楽情報変換手段3'により音楽情報へ変換され、音楽片生成手段32により音楽片へ分割されることもある。

【0030】上記処理を行った検索キーは、楽曲検索手段4へ送られ、楽曲検索手段4では、検索キーに類似する楽曲の一部もしくは全体をデータベースに蓄積された楽曲から検索する(ステップS22、S43)。データベースでインデックスが作成されている場合は、インデックスを用いて検索を行う。

【0031】音楽片を用いた検索では、音楽片検索手段

51により、検索キーとして送られてきた音楽片の少なくとも一つを用いて、データベースに蓄積された音楽片から類似するものを検索し、楽曲出力手段6により、検索結果の楽曲を出力する(ステップS23、S44)。

検索キーとして送られてきた音楽片の振動を用いて検索する場合には、検索キーのそれぞれの音楽片に類似するデータベース中の音楽片を検索した後、音楽片整合手段52により、音楽片同士の間隔性を調べ、最終的な検索結果とすることもある。

【0032】なお、音楽片の整合については、上記の開示の特許文献(特願平10-341754号)の明細書等に記載されている手法を採用する。具体的には、検索結果として検出された音楽片同士の関係と、検出された音楽片のそれぞれの検索キーとなった音楽片同士の関係を調べ、また、検出された音楽片の前後の音符と、検出された音楽片の検索キーとなった音楽片の前後の音符の関係を調べること、検索キーとなった楽曲の一部または全部と、検出された音楽片を含むデータベース中の楽曲の一部または全部との間の類似度を調べ、

【0033】楽曲入力手段1は、一定の時間間隔もしくは決められた規則に従った時間間隔で信号を送るリズム信号発信手段11を有することもある。特に、検索キーとして入力される楽曲に関しては、テンポ規格化が十分な精度で行われない可能性もある。そこで、リズム信号発信手段11によって一定の時間間隔もしくは決められた規則に従った時間間隔で信号を送信させ、その信号に合わせて楽曲入力者が楽曲の入力を行い、入力された楽曲と、リズム信号発信手段11により送信される信号とを同時にテンポ規格化手段2に入力させることによって、正確なテンポ規格化を行うことを可能とする。

【0034】あるいは、楽曲入力手段1は、リズム信号発信手段12を有することもある。これは、楽曲入力者が楽曲の演奏(歌唱)に合わせて発するリズム信号を受信可能とする手段で、入力された楽曲と、リズム信号発信手段12により受信される信号とを同時にテンポ規格化手段2に入力させることによって、正確なテンポ規格化を行うことを可能とする。

【0035】テンポ規格化手段2は、楽曲入力手段1により入力された楽曲、音楽情報変換手段3により変換された演奏情報、もしくは、リズム信号発信手段11により送信された信号、リズム信号受信手段12により受信された信号から、テンポを検出するテンポ抽出手段21を有することもある。この場合、テンポ規格化手段2は、テンポ抽出手段21によって抽出されたテンポを用いて、テンポ規格化を行う。

【0036】以下に、本発明の作用を説明する。図4は、テンポ規格化について説明するための図である。楽曲においては、テンポによって、音楽的な「拍子」の時間的な長さが固定される。図5(A)の「曲の楽曲」に示されているように、楽曲はそれぞれテンポを持ってあ

り、このテンポは、楽曲同士の間に一致することもあれば、異なることもある。また、同じ楽曲であっても、一つ一つの演奏によって異なることもある。また、同じ楽曲であっても、曲の部分によって異なることもある。

【0037】テンポを規格化するということは、図4(B)に示すように、すべての楽曲のすべての演奏の、すべての部分において、拍子の長さが一定の時間的な長さになるように、楽曲に変換を施すことである。

【0038】図5は、マッチングの単位を説明するための図である。テンポ規格化を行わない場合、主に音符がマッチングの単位とされてきた。しかし、図5(A)に示すように、音符の挿入・削除が生じると、比較する音符同士の対応関係がずれるという問題があった。D/Pマッチングは、音符の挿入・削除があった場合の比較する音符同士の対応関係のずれを修正するための方法である。

【0039】実際にD/Pマッチングを用いて検索を行う場合、許容する音符の挿入・削除の量には制限を設ける必要があるため、音符の挿入・削除が検索精度に大きな影響を与える。また、D/Pマッチングを用いた検索では、検索を高速に行うためのインデクスの作成などもできないため、多くの検索時間を必要とする。

【0040】一方、テンポ規格化を行った場合、図5(B)のように時間を単位としたマッチングが可能となる。この場合、音符の挿入・削除があったとしても、対応関係にずれが生じることがなく、そのことにより検索精度が影響されることはない。したがって、精度の高い検索が可能となる。また、テンポ規格化された楽曲を、音楽片に分割し、インデクスを作成することにより、検索を高速に行うようにすることも可能である。

【0041】本発明に関連する技術として、特願平10-329131号、特願平10-341754号に開示する説明がある。本発明は、データベース構築の際に、データベースに入力する楽曲のテンポ規格化を行い、検索キーとなる楽曲の入力の際もテンポ規格化を行う点で、先の発明と主に異なる。特に、本発明は、特願平10-341754号に開示する発明とは、特に音符単位での検索を行わず、時間を単位とした検索を行う点が主に異なる。

【0042】「発明の実施形態」(第1の実施形態)本発明の第1の実施形態について説明する。データベースに入力するための楽曲として、MIDIデータを用いる。現在ではカラオケ用のデータを加え、多くの楽曲がMIDIデータとして世の中に出回っている。

【0043】MIDIデータでは、音階情報が必ず含まれると規定されている。そして、拍子の時間的な長さが、テンポ情報として、音階情報とは別に記述されている。したがって、規格化された時間として、拍子を基準とした時間を採用すれば、データベース中の楽曲のテ

ンポ規格化が実現される。

【0044】データベースに格納するにあたり、元のMIDIデータは、まず、音楽片に分割される。分割の方法としては、音楽片の長さが、規格化された時間軸上ですべて同じになるようにし、図8(A)に示すように、すべての音符の始まりを先頭とする音楽片が生成されるようにする方法がある。あるいは、図8(B)に示すように、音楽片が一定の時間間隔で始まるように分割する方法もある。

【0045】また、音楽片へ分割した後、音楽情報(特徴量)として、音高推移ベクトルを生成する。これは、図7に示すように、規格化された時間軸を等間隔の目盛で読み、各目盛の位置で演奏されている音の高さを並べてできるベクトルである。計算機での処理上、音の高さは、例えばMIDIに使用されている音高を数値や、周波数などに置き換える。

【0046】音楽情報(特徴量)としては、上記のほかにも、例えば、以下のものを生成することがある。

【0047】① 音高推移ベクトルA:  $= (a_1, a_2, \dots, a_n)$  としたとき、ベクトルの要素の平均値  $m = (a_1 + a_2 + \dots + a_n) / n$  を計算し、各要素からこの平均値を引いた値を要素とするベクトルA:  $A' = (a_1 - m, a_2 - m, \dots, a_n - m)$  を特徴量とする。検索キーの楽曲と、データベース中の楽曲の調が異なっても、このベクトルは類似することが期待される。

【0048】② 音高推移ベクトルA:  $= (a_1, a_2, \dots, a_n)$  に対し、最初の要素  $a_1$  を基準にし、各要素からこの基準値を引いた値を要素とするベクトルA:  $A' = (0, a_2 - a_1, \dots, a_n - a_1)$  を特徴量とする。検索キーの楽曲と、データベース中の楽曲の調が異なっても、このベクトルは類似することが期待される。

【0049】③ 音高推移ベクトルA:  $= (a_1, a_2, \dots, a_n)$  に対し、各要素から前の要素の値を引いた値を要素とするベクトルA:  $A' = (a_2 - a_1, a_3 - a_2, \dots, a_n - a_{n-1})$  を特徴量とする。検索キーの楽曲の調が少しずつ推移してしまっても、データベース中の楽曲と、このベクトルは類似することが期待される。

【0050】これらの音楽情報が、データベースに格納された後、検索を高次元するためのインデックスを生成する。インデックスの生成方法として、K-tree などを用いる。なお、K-tree に関する参考文献としては以下のものがある。

【参考文献】S. Gutfreund, "Efficient dynamic index structures for spatial searching", *Information Systems for the 21st Century International Conference on Management of Data*, pp. 47-57, Boston, 1994.

検索キーとなる楽曲の入力については、ハミングによる

入力、楽器の演奏による入力、MIDIキーボードの演奏による入力などの方法がある。

【0051】楽曲の入力の際には、リズム信号受信装置として、電子的なメトロノームを用いる。この電子的なメトロノームは、指定したテンポで音や光、コンピュータディスプレイ上の文や、図形などを発することができるようになっており、楽曲入力者が、その発する信号に合わせて楽曲の入力ができるようになっている。この電子的なメトロノームは、一定の時間間隔で信号を発するだけでなく、例えば、小節の先頭の拍は青色・音響や、光の色・光の強さ、ディスプレイ上の文字の種類・図形の種類の、他の拍とは異なるように発したり、また例えば「サンバのリズム」など、入力する曲の種類に合わせてリズム信号を発することができ、また、この信号を発するテンポであるが、楽曲入力者が楽曲の入力をしやすいように、遅くしたり速くしたりを自由に設定することが可能である。

【0052】あるいは、楽曲入力の際には、リズム信号受信装置を使うこともある。これは、手拍子、足踏み、指の上下の動き、体の揺れ、マウスのクリック、キーボードの押下など、楽曲入力者が楽曲入力の際に意識的／無意識的に発する動き、信号などを検知できるようにした装置である。

【0053】入力された楽曲と、上記リズム信号受信装置から得られる信号を使って、検索キーとなる楽曲のテンポ規格化を行う。

【0054】音楽片への分割、音楽情報の変換については、MIDIキーボードから入力される楽曲については、データベースへの楽曲の入力の場合と同じである。

【0055】ハミングによる入力、楽器の演奏による入力では、楽曲入力手段1によりPCM形式のデータとして楽曲が入力される。そして、このデータから高品質の抽出される。音楽情報の抽出に際しては、PCM形式のデータから音高情報を抽出する従来の方法に準じる。

【参考文献】新藤高水、今村正和、井口健士、「聴覚の自動認識」、計測自動制御学会論文集、Vol.20, No.10, pp.949-945, 1984; 【参考文献3】水戸正典、藤本正樹、高島幸雄、鶴田七郎、「パーソナルコンピュータシステム 一歌声の自動認識」、情報処理学会第35回全国大会、5F-5, 1997

すなわち、まず、FFTや、自己相関処理を用いて周波数分析を行い、ピーク抽出、各ピークのパワー測定などを行う。さらに、ノイズの除去、ビッチドリフトの補正、高調波成分の基本成分への取扱いなどを行い、時間軸上の目盛に対して、どの時点でどの音が演奏されているかを検出する。音楽情報が抽出された後の、音楽片への分割、音楽情報の生成の仕方については、データベース中の楽曲の入力の場合と同じである。

【0056】なお、音楽情報生成の過程を二段階に分け、途中でテンポ規格化を行う場合もある。例えば、引

11

ず特徴情報を抽出した後、テンソルを規格化した後、音高ベクトルなど特徴量の生成を行う場合もある。

【0057】検索については、まず、検索キーとして入力された楽曲から生成された音楽片のそれぞれに類似する音楽片を、データベース中から検索する（音楽片検索）。その際には、検索キーの音楽片とデータベース中の音楽片との距離を、音高ベクトルなどそれぞれの特徴量に対して求める。なお、この距離は、ユークリッド距離やマンハッタン距離など適当な距離を特徴量ごとに設定して求める。また、検索を高速化するため、インデクスを用いて検索を行う。

【0058】音楽片同士の距離（非類似度）は、各特徴量の距離に適当な重みを加味して何らかの方法で求めたものとして表す。詳しくは、以下のとおりである。用いている特徴量の集合を $FSPACE$ とする。音楽片 $K$ と音楽片 $X$ との特徴量 $i$ についての距離を $d_i(K, X)$ 、重みを $W_i$ とすると、音楽片 $K$ と音楽片 $X$ との総合的な距離 $D(K, X)$ を何らかの方法で求めるのは、次のような関数 $f$ を用いて行うことになる。

【0059】

$$D(X, K) = f(\{W_i\}, \{d_i(K, X)\})$$

（ただし、 $i \in FSPACE$ ）

この関数 $f$ の具体的な例として、線形和の関数を用いると、

$$D(X, K) = \text{sum}(W_i \times d_i(K, X))$$

となる（ただし、 $i \in FSPACE$ であり、 $\text{sum}$ は $FSPACE$ についての総和をとる）。

【0060】関数 $f$ として線形和ではなく、例えば次のような最小値をとる関数を用いることもできる。

【0061】

$$D(X, K) = \min(W_i \times d_i(K, X))$$

（ただし、 $i \in FSPACE$ 、 $\min$ は $FSPACE$ についての最小値をとる）

また、上記線形和と最小値との両方を用いることもできる。すなわち、検索用 $FSPACE$ を2種類（ $FSPACE1$ 、 $FSPACE2$ ）に分けて、 $FSPACE = FSPACE1 \cup FSPACE2$ であるとすると、

$$D(X, K) = \text{sum}(W_i \times d_i(K, X)) \cdot \min(W_j \times d_j(K, X))$$

のような関数を用いて総合距離を求めることもできる。ここで、 $i \in FSPACE1$ 、 $j \in FSPACE2$ である。

【0062】ここでは、音楽片同士の距離（非類似度）の求め方として、線形和を用いる場合、最小値を用いる場合、線形和と最小値の両方を用いる場合の例を説明したが、もちろんこの他に非類似度の求め方はいろいろある。データベース中の音楽片のうち、検索キーの音楽片との距離が近いもの、すなわち類似度の高いものが音楽片検索の結果となる。

12

【0063】検索キーから音楽片が複数生成される場合、検索キーの音楽片のそれぞれに対する音楽片検索の結果に対して音楽片同士の整合性を測り、複数の音楽片を組み合わせた場合の類似度を計算し、最終的な検索結果とすることができる。

【0064】今、検索キー $K$ の音楽片 $K_1, K_2, \dots, K_n$ （ $p < q$ ）に対し、データベース中の音楽片 $X_1, X_2, \dots, X_m$ がそれぞれ類似していたとする（ $K_1 \sim X_1, K_2 \sim X_2, \dots, K_n \sim X_n$ ）。検索キーの音楽片 $K_1, K_2, \dots, K_n$ に対して整合性を持つということは、以下の条件を満たすことである。

【0065】1) 音楽片 $X_1, X_2, \dots, X_m$ は同じ曲に属さなければならない。なお、以降ではこの候補曲を $A$ とし、音楽片 $X_1, X_2, \dots, X_m$ をそれぞれ $A_1, A_2, \dots, A_m$ とおく。

【0066】2) 音楽片 $A_1, A_2, \dots, A_m$ の時間的順序は、対応する検索キーの音楽片（ $K_1, K_2, \dots, K_n$ ）の時間的順序と同じ（すなわち、 $r < s$ ）である。

【0067】3) 音楽片 $A_1, A_2, \dots, A_m$ の規格化された時間における開始時間を、 $T_{10}, T_{11}, \dots, T_{1m}$ とし、検索キーの音楽片 $K_1, K_2, \dots, K_n$ の規格化された時間における開始時間を、 $T_{20}, T_{21}, \dots, T_{2n}$ としたとき、開始時間の間隔が一致、すなわち、 $T_{10} - T_{11} = T_{20} - T_{21}, \dots, T_{1m-1} - T_{1m} = T_{2n-1} - T_{2n}$ である。

【0068】検索結果の出力では、検索キーと、検索結果の候補となったデータベース中の楽曲の一部もしくは全体との間の類似度を求めた後、類似度の高い順番に、楽曲の一部もしくは全体を提示する。データベース中の情報から楽曲の題名、音楽家名を表示することもできるし、検索キーと類似する楽曲の一部もしくは全体を演奏することもできる。

【0069】本実施の形態においては、データベースに入力する楽曲としてMIDIデータを用いることかできる。MIDIデータは標準化された形式で、非常に多くの楽曲がMIDIデータとして存在しており、データベースの構築のために楽曲の演奏を行ったりする必要がなく、データベース構築が容易である。また、MIDIデータは拍子を基準として音符情報で記述されているため、テンソル抽出、特徴情報の抽出が正確に行えるため、音楽検索用のデータベースに利用することに適している。

【0070】また、検索キーとなる楽曲の入力の際に、リズム（打点位置）情報も自動的に使うことによって、検索キーとなる楽曲のテンソル規格化が正確に行えるようになる。

【0071】本実施の形態の他の効果として、データベースの検索が容易かつ正確に行えること、また、検索キーとなる楽曲のテンソル規格化が正確に行えることが挙げられる。

【0072】（第2の実施の形態）本実施の第2の実施の形態においては、上記時間1の楽曲の演奏でデータベースに入力する楽曲としてMIDIデータを用いるのではなく、例えば音楽用CD（コンパクト・ディスク）の



## 13

データや、WAVE、AIFC、AUなどの形式のよう  
にPCM符号化された楽曲。もしくは、他の符号化形式  
のものであれば、それを一度PCM符号化形式に変換し  
たものを用いる。

【0073】一般に楽曲においては、拍子に合わせて音  
が演奏される。そして、拍は、演奏が開始される瞬間に  
大きく音圧が変化するため、楽曲における音圧の変動を  
調べれば、テンポ情報を抽出することが可能である。

【0074】より具体的なテンポ情報抽出方法として、  
以下のような方法がある。

【0075】① 100Hzや10Hzといった小さな  
サンプリング周波数で音圧レベルのピークを探し、ピー  
クの間隔からテンポを抽出する方法。元のPCMデータ  
のサンプリング周波数が、例えば44100Hzであつ  
たとすると、1000サンプルずつ固定したあわせたデータ  
を作成し、音量レベルのピークを探し、といった方法が  
考えられる。

【0076】② 上記小さなサンプリング周波数で音量  
レベルの変化率の推移を調べ、変化率が急激に増加する  
部分を探し、その直後に増加する部分の間隔からテンポ  
を抽出する方法。

【0077】③ 上記小さなサンプリング周波数のデー  
タに対し、例えば、10秒といった長い区間で周波数分  
析を行い、ピークを探索する方法。例えば、2秒のところ  
にピークが出れば、この楽曲の1拍の長さは0.5秒で  
あるということになる。

【0078】④ リズムを刻むドラムズなどの楽器の音  
を選択的に抽出し、これらの音の音圧の変動を分析する  
ことにより、テンポ情報を抽出する方法。

【0079】入力された楽曲から音楽情報などの特徴量  
を生成する方法としては、上記の第1の実施形態でハ  
ミングや楽器の演奏により入力された楽曲から音楽情報  
などを抽出する方法に準じる。ただし、ハミングや単一  
の楽器の演奏の場合と違い、音源用CDなどのデータで  
は、複数の音符が同時に発せられているため、検出対象  
となる音高（振動数）に相当する音の音高のみを抽出する  
必要がある。そのために、F2やF3と呼ばれる周波数帯域  
のピークは、音圧の最も強いピークを抽出する方法や、発声  
やリズムを演奏する楽器、もしくは人の歌の音域や音高  
のピークのみを抽出するような帯域フィルタを使用する  
方法がある。

【0080】本実施形態の第1の実施形態の第1の実施  
形態、データベースに入力する楽曲として、第1のより  
さらに一般的な音楽符号化方式であるPCM方式により  
符号化された楽曲を利用することにより、データベース  
中に含まれる楽曲の量を、より増加したものにすること  
が可能である。また、市販の音楽用CDのみならず、個人  
的に音楽によって制作した楽曲などもデータベース中  
に含まれることができるようになる。

【0081】第1の実施形態の第1の実施形態の第1の実施

## 14

の形態においては、上記の第1、第2の実施形態でデ  
ータベース中の楽曲、検索キーの楽曲を音楽片に分割  
し、特徴量を生成する際に、音楽情報ベクトルを生成す  
るのではなく、対象となる音楽片における音高の分布を  
求め、この分布をもとに特徴量ベクトルを生成する。

【0082】ある音楽片において、音高が、図8(A)  
に示すように推移した場合、規格化された時間の1単位  
を一度に計算して、図8(B)のような音高分布図を  
作成し、各音高（もしくは音高の範囲）の度数を並べ  
て音高分布ベクトルを作成することができる。これを、音  
楽情報（特徴量）として検索を行うことができる。

【0083】音楽情報（特徴量）としては、上記のほか  
に、例えば、以下のものを生成することができる。

【0084】① 第1の実施形態で利用した音高推移  
ベクトルA、の各要素から、要素の平均値を引いてでき  
たベクトルA、'の要素の値の分布。

【0085】② 第1の実施形態で利用した音高推移  
ベクトルA、の各要素から、最初の要素の値を引いてでき  
たベクトルA、'の要素の値の分布。

【0086】③ 第1の実施形態で利用した音高推移  
ベクトルA、の各要素から、前の要素の値を引いてでき  
たベクトルA、'の要素の値の分布。

【0087】また、上記各分布において、図8(B)の  
音高分布図の横軸に相当する値が0に相当する部分の度  
数を0とする場合もある。また、上記各分布において、  
図8(B)の音高分布図の横軸に相当する値を、定めら  
れた範囲内に限定し、範囲を超えたデータを削除する場  
合、もしくは、その範囲の境界の値に加える場合もある。

【0088】これらの特徴量ベクトルを利用し、第1の  
実施形態の方法に従い、音楽片検索、音楽片整合を結  
した上で類似検索を行う。

【0089】本実施形態の第2の実施形態について、第1の  
実施形態の場合よりも、一般に特徴量ベクトルの次元数を  
減らすことができるために、検索精度を向上させることが  
可能である。また、音楽片の範囲内では、時間の情報  
が失われるために、データベース中の音楽片と、検索キー  
から生成された音楽片の開始地点に多少の相違があつ  
ても、検索が可能となる。一方、時間の情報が失われる  
ために、音高推移ベクトルで見れば類似しないもので  
も、音高分布ベクトルでは類似する可能性がある。した  
って、まず本実施形態の第2の実施形態により検索を行  
い、検索結果の中から、第1の実施形態による検索を行  
うよりさらに検索などを行うことも可能である。

【0090】【類似の楽曲】上述のように、本実施形態では、検索  
キーがハミングや楽器の演奏などにより、自由に楽曲の一  
部分または全体を指定することにより、検索者が指定した  
楽曲やデータベースの中から検索より迅速に検索すること  
が可能となる。

【0091】本発明によって、例えば、カラオケや、ミュージック・オン・デマンドシステム（自動ジョークボックス）などのシステムにおいて、利用者が選択したい楽曲の名前や作曲者名、演奏家名などを思い出せない場合においても、その楽曲のメロディーさえ覚えていれば目的の楽曲を選択可能となり、このような分野において極めて有効である。

【図面の簡単な説明】

【図1】本発明による音楽の類似検索（音楽情報への変換を行わない場合）の概要を説明する図である。

【図2】本発明による音楽の類似検索（音楽情報への変換を行う場合）の概要を説明する図である。

【図3】本発明の原理構成図である。

【図4】テンポ正規化を説明する図である。

【図5】マッチングの単位を説明する図である。

【図6】音楽片への分割を説明する図である。

【図7】音楽情報への変換を説明する図である。

【図8】音楽情報への変換（音高分布）を説明する図である。

【図9】従来の方法による音楽検索（Dマッチング使用）の原理構成図である。

\* 【図10】本発明に関連する音楽片を使用する音楽検索の原理説明図である。

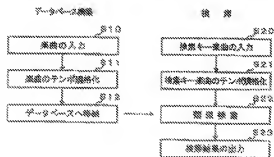
【図11】本発明に関連する音質情報を使用しない音楽検索の原理説明図である。

【符号の説明】

- 1 楽曲入力手段
- 11 リズム信号受信手段
- 12 リズム信号受信手段
- 2 テンポ正規化手段
- 10 21 テンポ抽出手段
- 3 1 音楽情報変換手段
- 31 音楽情報変換手段
- 32 音楽片生成手段
- 4 楽曲部群手段
- 41 楽曲部群部
- 42 音楽情報部群部
- 43 音楽片部群部
- 44 インデックス生成手段
- 5 楽曲検索手段
- 20 51 音楽片検索手段
- \* 52 音楽片整合手段

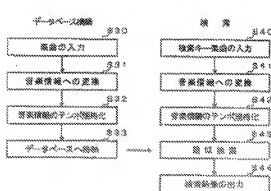
【図1】

本発明による音楽の類似検索（音楽情報への変換を行わない場合）の概要



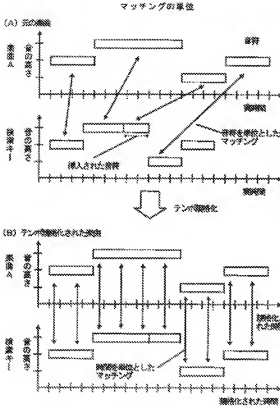
【図2】

本発明による音楽の類似検索（音楽情報への変換を行う場合）の概要

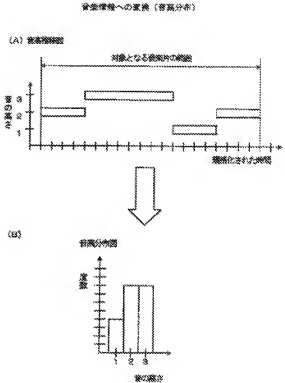




【図5】

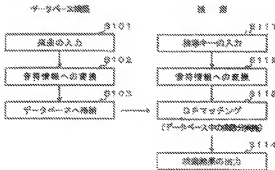


【図6】



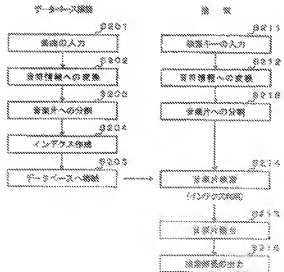
【図9】

従来の方法による音楽検索（コママッピング使用）の処理フロー図



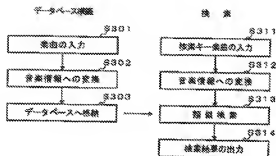
【図10】

音楽片を使用する音楽検索の音楽情報図



【図11】

貸行情報を使用しない音楽検索の処理フロー図



フロントページの続き

(72)発明者 坂田 哲夫  
東京都千代田区大手町二丁目3番1号 日  
本電信電話株式会社内

(72)発明者 山室 雅明  
東京都千代田区大手町二丁目3番1号 日  
本電信電話株式会社内

(72)発明者 梅田 昌茂  
東京都千代田区大手町二丁目3番1号 日  
本電信電話株式会社内

(72)発明者 紺谷 精一  
東京都千代田区大手町二丁目3番1号 日  
本電信電話株式会社内